

SOME PSYCHOLINGUISTS' REACTIONS TO THE SYMPOSIUM ON DYNAMIC ASPECTS
OF SPEECH PERCEPTION

W.J.M. LEVELT and G.B. FLORES D'ARCAIS

The organizers of this Symposium have asked us to give some psycholinguists' reactions to what has been reported and discussed during this conference. They presumably hoped that such naive observers would be able to look objectively at the emperor's clothes. Our reaction to this is two-fold: (1) Though it may be disappointing, we have not observed any nudity. On the contrary, we were impressed by the richness of methods, phenomena and theoretical concerns among speech perception (SP) researchers. (2) We did not feel excentric to this meeting at all. Not only are many speech researchers psychologists by training, but also more generally we experienced considerable openness with respect to psychological theory, and psycholinguistic experimentation in language perception. If one compares the present situation with earlier work in SP, one finds a strongly increased interest in syntactic, semantic and even textual aspects of the speech signal. There is certainly a realistic common base now for students of "speech" and "language" perception to cooperate, and to integrate their theories. It even seems to be the case that the vigorous appeal of SP-workers to psychological models of memory, especially the early storage mechanism, is forcing psychologists into granting that their models of handling the intake and storage of continuous speech are still inadequate.

Since many of the observations we made have already been brought up by others (see especially Pisoni's paper) it will suffice for us to mention one or two issues which, in our view, still have not received sufficient attention.

- In spite of the just-mentioned convergence in the work of speech researchers and psycholinguists, there are still remarkable differences in style of theorizing. Though there are exceptions on both sides, SP-theories seem to be more structure-oriented, whereas psycholinguists show a deeper interest in process. Many of the SP-studies at this Symposium were of the following sort: here is a variable (linguistic, acoustic, etc.), does it affect a subject's reactions? Examples: Does the number of syllables surrounding a vowel affect the subjects' identification of a

vowel? Can lip-rounding be detected in vowel-preceding consonant clusters? Is voice onset time "real", i.e. influencing phoneme identification? Are syntactic variables "real" in judging segment duration? Does F_0 affect a subject's judgment with respect to the position of a sentence in a paragraph? Do intonation and syntax affect judgments of dichotic channel-switching? Though this type of question often leads to quite unexpected and non-trivial results it reminds us too much of the so-called "psychological reality" studies of the sixties. Psycholinguists have mostly passed that state now. Their interest is in the processes underlying such findings: what strategies and procedures are followed by the subject to cope with the experimental task? Which are the successive stages in the subject's decision making etc. Usually, this is called information processing approach, and we have seen too little of that during this conference. There is, for instance, not a single flow-diagram (i.e. a visual representation of a sequence of operations and decisions) in any of the Symposium papers. (There were some block-diagrams, but they are still rather "structuralist" in that the exact nature of the procedures within the blocks is left unspecified). Of course, one can do without flow-diagrams, but in our opinion one cannot do without information processing models in SP-research.

- A related point is that we have seen little concern for task structures. As in the early psychological reality studies, the objective of many SP-studies seems to be to demonstrate that a certain independent variable is effective, without much concern for the experimental task. However, since the work of Newell and Simon we know that the structure of the task is the prime determinant of the subject's decision procedure, and that there are only a few general invariant characteristics in human information processing (such as capacity-limits). Many of the SP-studies of this Symposium could have profited from a thorough task analysis: what exactly is the subject asked to do, what are the possible ways for him to cope with that task, and can we empirically determine which of these procedures is in fact chosen by him? Such an analysis might add considerably to an understanding of various interesting phenomena in SP (like those reported by Wingfield), where at present we do not even know whether they arise in early

stages of stimulus encoding, or in later phases of response selection. Also, it is not unlikely that much of the rather "structuralist" controversies about "fixed" structures such as early storage systems will disappear if more attention is given to finer details of the task structure.

Finally, highly interesting theoretical issues will come up: given a certain task and the possible alternative ways to cope with it, why is it that a subject selects one procedure rather than another to arrive at an adequate response?

- An important point which has emerged several times in some of the papers presented and in the discussion, is the issue of parallel versus serial processing in speech perception. The question has been put forward by Haggard in his stimulating introduction, and has been given close attention in the last paper of the Symposium by Pisoni.

When one looks at the various models of speech perception available in the literature, it clearly appears that they are essentially, at least in most stages postulated, serial. Most models assume first some "low level" acoustic processing or preprocessing, which produces on the incoming signal certain "corrections" and equalizations. The material is then held for a brief duration in some storage of a sensory type, where some integration may take place, and then is handled over to some kind of perceptual unit where real "perceptual processing" takes place. Notice that, while some parallel processing may take place at the very first stage of acoustic processing, processing in the later stages, at least until the material reaches the level of LTM, essentially takes place sequentially. As a consequence, in most models real semantic processing would start only at a later stage. Now, it may be true that in certain experimental tasks no semantic processing takes place (for example, in a situation where the S has to search for a certain acoustic target in a series of discrete acoustic events presented in isolation) However, there is an increasing amount of evidence, even among data reported in the discussion during the Symposium, that in perception of connected speech processing takes place simultaneously in parallel, and this, very likely, from very early stages. In the attempts toward a construction of a speech perception theory, we wonder whether more attention should not be given to the possibility of very early semantic and syntactic processing, on the basis of linguistic or even non-linguistic context, of the expectations of the

- hearer etc. which could contribute to impose certain restrictions as to the kind of signals which may be accepted as possible candidates for perceptual synthesis, or produce certain selections or modifications on the incoming acoustic events. We do not propose to put a lexicon already on the basilar membrane, we only want to emphasize the notion that from the very early stages processing may take place at several levels simultaneously, and that the perceptual result may be the output of alternative or simultaneous activation of one or more of such levels of processing. There is a substantial amount of evidence from visual information processing studies which seems to indicate that parallel processing models can explain some critical phenomena in visual search paradigms in a more appropriate way than serial models. We would like to see more attention for a similar idea for acoustic information processing in speech perception models.
- A question which is related is whether the sensory storage which is postulated, with different labels, as an essential component in all models of speech perception, and in which the incoming information is held for a brief interval, during which some integration may take place, is to be considered a kind of "buffer" in which no recoding takes place, or whether some processing or some selection may already be at work. and that a certain amount of recoding may take place at this level. In the model proposed by Pisoni at this seminar, the S.I.S. component is functionally connected with LTM, which may mean a certain amount of interaction between SIS and LTM; in the pre-perceptual storage proposed by Massaro, on the other hand, no such connections are postulated. The question whether at the sensory mode level some recoding has already taken place, what kind of recoding is possible there, and under the influence of which factors is still, in our opinion, an open question, and specific answers in this direction are no doubt of importance in the development of a theory of speech perception.
 - Another issue of major importance seems to us the series of questions one may ask about the effect of context on the perception of speech events. Various papers paid attention to several contexts, and when one compares the approach taken with the one typical of less recent work on speech perception, which was mainly concerned with very short units such as vowels or CV combinations out of any context, it seems already a real progress to have taken into account larger linguistic contexts such as words or sentences. This may have increased the ecological representativeness of the studies, and cer-

tainly may have contributed toward a theory of speech perception less limited in its explanatory, descriptive and predictive power. Recent psycholinguistic work is taking into account larger and larger contexts. Serious attention is given not only to the larger linguistic context of a sentence, but also to the non-linguistic context, that is, to the different aspects of the communication situation - including things like the "knowledge of the world" of the two partners in the communication act, of the possible referents present, of events of which both the partners are aware, etc. Briefly, to what are often called presuppositions of a linguistic event. Several studies have shown that such presuppositions do affect perception, comprehension, storage and retrieval of linguistic material. and it is obviously an interesting question to ask, to what extent such presuppositions may affect perception of so-called "simple" linguistic events such as phonemes. We would welcome an effort in this direction.

- A question which is implicit in all theorizing about speech perception, and which has been specifically treated in Prof. Öhman's paper, is about the definition of perception. We do not think that it is possible to discuss all complex problems connected with such a definition in few words. However, a couple of comments may be appropriate here.

First, we must be aware of the different usages of the word in different experiments and in theory construction. In speech research the word perception has been used in a variety of ways. In some cases the experimenter means by perception the sensory discrimination between, say, two single acoustic events of different duration, and the paradigm adopted is the one typical of psychophysical experiments. In other cases the term is used to indicate the complex process of understanding the meaning of a word in a given context. While this variety of uses is characteristic of other fields of perception as well, one must be very alert, here as in other fields, to possible traps. An important point here is the opportunity of not confusing situations which call upon judgments on a certain speech event, with "real" perceptual situations. The interpretation of certain "click" experiments is an example of this confusion, where effects, which can clearly be assigned to response bias, are interpreted as

perceptual effects.

- Another point which often has emerged in the discussion is the problem of the "perceptual unit" in speech perception. The question is whether it makes sense to talk about such units and state, for example, that "the syllable is the basic unit of speech perception." In visual perception nobody has ever found it important to decide which is the "basic perceptual unit", and research has progressed anyhow. The question could perhaps be reformulated in terms like the following;
 - a. How is the speech event segmented and interpreted?
 - b. Through what processing strategies do we obtain a certain perceptual result?
 - c. If we want to isolate certain units such as the syllable or the word, do we know how these units are organized within the perceptual events?

- A final comment relates to the relevance of developmental studies for a theory of speech perception. There was only one developmental paper at the Symposium, but there are good reasons for SP researchers to give more attention to children's speech. Firstly, in the present literature developmental studies are rather more concerned with speech production than perception; more perceptual work is needed to achieve general understanding of speech development. Secondly, SP-studies with children may contribute to a general theory of speech perception. Specific theoretical issues where such contribution can be made are among others:
 - (1) the question whether findings in SP should be interpreted as specific to language, or as special cases of more general perceptual abilities (cf. Eimas' studies relating categorical perception to neonate speech discrimination).
 - (2) questions concerning the hierarchy of perceptual difficulty of various aspects of the speech signal, and concerning the hierarchy of SP skills. The order of acquisition of perceptual skills in children may not only reflect the order of complexity of perceptual tasks, but also the hierarchical skill structure of the adult. Early acquired perceptual strategies may remain basic strategies in adulthood.Furthermore, the child's speech perception may be used methodologically to test certain constraints in SP-theory. For instance,

parameters for certain hypothesized types of storage, will probably be different for the child, so that the functional contribution of such mechanisms to SP can be evaluated by comparing the child's and adult's behavior with respect to identical perceptual tasks.